Automatic detection of clustered microcalcifications using a combined method and an SVM classifier.

Armando Bazzani [1,3], Alessandro Bevilacqua [2,3], Dante Bollini [1,3], Renato Campanini [1,3],

Nico Lanconelli [3], Alessandro Riccardi [1,3], Davide Romani [3]

[1] Department of Physics, University of Bologna, Italy

[2] Department of Electronics, Computer Science and Systems, University of Bologna, Italy

[3] National Institute for Nuclear Physics, Bologna, Italy

Abstract. In this paper we investigate the performance of a Computer Aided Diagnosis (CAD) system for the detection of clustered microcalcifications in mammograms. Our detection algorithm consists on the combination of two different methods. The first one, based on filtering techniques and gaussianity statistical tests, finds out the most obvious signals. The second one is able to discover more subtle microcalcifications by exploiting a multiresolution analysis by means of the wavelet transform. In the false-positive reduction step we separate false signals from microcalcifications by means of a Support Vector Machine (SVM) classifier.

Our algorithm yields a sensitivity of 95% with 0.6 false positive cluster per image on the 40 images of the Nijmegen database.

1. Introduction

Breast cancer is the most common form of cancer among women. The presence of microcalcifications in breast tissues is one of the main features considered by radiologists for its diagnosis. Several techniques developed for the automated detection of microcalcifications can mainly be grouped in three different categories: multiresolution analyses (Yoshida et al. 1994; Lado et al. 1999), difference-image techniques (Chan et al. 1987) and statistical

methods (Karssemeijer 1993; Gürcan et al.1998). By comparing the different methods it turns out that some microcalcifications are detected by one method but missed by others.

In this paper we propose an approach based on the combination of different detection methods in order to get optimal performances. Yoshida *et al* pointed out that the simultaneous use of two or more techniques might improve the results of an optimized single method (Yoshida et al. 1996). The basic idea of our method is to combine a multiresolution analysis based on wavelet transform with a filtering method and a gaussianity statistical test and to perform a logical OR operation on the detected microcalcifications before clustering.

In the false-positive reduction (*fpr*) step we try to separate false signals from microcalcifications by using a classifier based on a Support Vector Machine (SVM).

The detection scheme has been tested on 40 digitized mammograms coming from Nijmegen Hospital.

2.   Methods

2.1. Detection scheme

Microcalcifications are very small spots that are relatively bright compared with the surrounding normal tissue. Typically they are between 0.1 mm and 1 mm in size and are of particular clinical significance when found in clusters of five or more in a 1 cm^2 area. Most of the clusters consist of at least one evident microcalcification and other more hidden signals.

Our approach includes two different methods: the first one (coarse) is able to detect the most obvious signals and uses filtering techniques and gaussianity tests, while the second one (fine), based on multiresolution analyses, discovers more subtle microcalcifications.

First the digitized image is segmented to isolate breast tissues from image background. In this way we reduce both the processing time and memory requirements, since we analyze only

areas which contain useful information for the detection. The segmented image is then passed to the two signal-extraction methods described in the following subsections. Signals coming out from these methods are combined through a logical OR operation and then clusterized to give the final result. The detection scheme of the algorithm is shown in Figure 1.
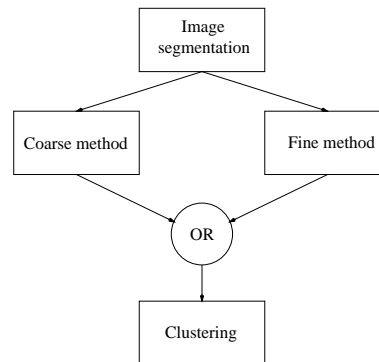

Figure 1. Detection scheme.

2.2. Coarse method

In this part of the algorithm we remove structured image background by means of a filtering technique. The scheme of the coarse method is shown in figure 2.
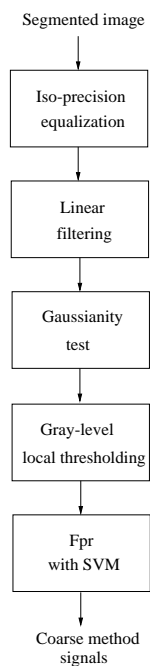

Figure 2. Scheme of the coarse method.

First of all we perform an iso-precision noise equalization as described in Karssemeijer 1993. The equalized image is passed through a linear filter:

$$x'_{i,j} = \frac{1}{(2N_1+1)^2} \sum_{n=-N_1}^{N_1} \sum_{m=-N_1}^{N_1} g1_{n,m} \cdot x_{i+n,j+m} - \frac{1}{(2N_2+1)^2} \sum_{n=-N_2}^{N_2} \sum_{m=-N_2}^{N_2} g2_{n,m} \cdot x_{i+n,j+m},$$

where $(2N_1+1)$ and $(2N_2+1)$ are the sides of the masks $g1$ and $g2$, $x_{i,j}$ and $x'_{i,j}$ are the gray values of the pixel (i,j) before and after filtering. $g1$ and $g2$ have the form:

$$g1 = \begin{bmatrix} .75 & .75 & .75 \\ .75 & 1 & .75 \\ .75 & .75 & .75 \end{bmatrix} \quad g2 = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}$$

According to experimental evidences we assume that the remaining noise is gaussian, since we have reduced the structured noise in the filtering step. We then employ a gaussianity test on the filtered image in order to choose ROI's that include interesting signals. Since this image contains only gaussian noise and signals with a high contrast we should have a deviation from gaussianity in regions including microcalcifications. Here we perform the gray-level local thresholding: the central pixel of the considered 51x51 window of the filtered image is retained only if its gray level is greater than the mean pixel value multiplied by a preselected k multiple of standard deviations σ.

The next step is a false-positive reduction (*fpr*) phase based on a local edge-gradient analysis (Ema et al. 1995). We consider five features (*area*, *average pixel value*, *edge gradient*, *degree of linearity*, *average local gradient*) to separate microcalcifications from false signals. These features are the inputs of an SVM classifier described in subsection 2.4.

Signals survived to the *fpr* stage will join others coming from the fine method described in the next subsection.

2.3. Fine method

In this part of the detection scheme we try to discover more subtle microcalcifications. In figure 3 the scheme of the fine algorithm is depicted. Microcalcifications are characterized by well-defined range size and high local contrast, so we find out signals having these features. We split the algorithm into two independent sections.
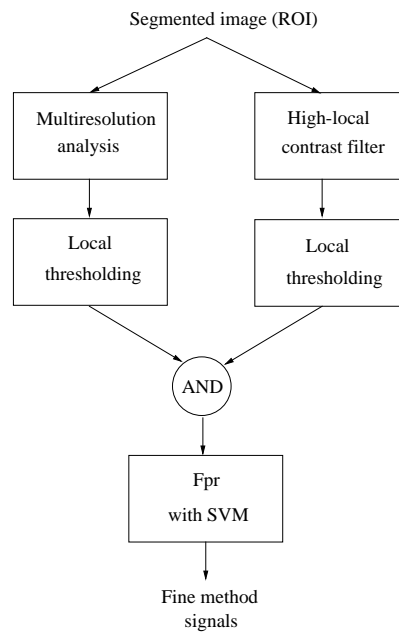
Segmented image (ROI)

Multiresolution analysis

High-local contrast filter

Local thresholding

Local thresholding

AND

Fpr with SVM

Fine method signals

Figure 3. Scheme of the fine method.

The first one detects signals having size smaller than 1 mm by means of a multiresolution analysis based on wavelet transform. To extract interesting signals we perform a local thresholding in 40x40 pixels size windows. Assuming a gaussian distribution for the noise we fit with a parabola the gray level histogram of the window in semi-logarithmic scale. Then we retain pixels having a gray level greater than the one intersecting the parabola and the x axis. Signals having a high local contrast are enhanced in the second section, by using a filtering technique. We subtract the image obtained by a 9x9 moving average filtering from the enhanced image coming from a 3x3 gaussian filter. We carry out the same local thresholding on the filtered image, followed by a morphological opening operation.

After that a logical AND operation is accomplished on signals extracted by the two sections of the fine method. To split false signals from microcalcifications a *fpr* phase similar to the one used in the coarse method is performed.

Finally, microcalcifications which have passed the *fpr* step are joined with others coming from the coarse method through the logical OR operator.

2.4. SVM classifier

The false-positive reduction step is a two classes pattern recognition problem: we must distinguish true microcalcifications from false signals. SVM are learning machines used in pattern recognition and regression estimation problems (Vapnik 1995). They grow up from Statistical Learning Theory (SLT) problems (Vapnik 1995), which gives some useful bounds on the generalization capacity of machines for learning tasks. The SVM algorithm constructs a separating hypersurface in the input space. It acts in this way:

a) maps the input space into a high dimensional features space through some non linear mapping chosen a priori (*kernel*);

b) constructs the Maximal Margin Hyperplane in this features space.

Hyperplanes are defined by $\mathbf{w} \cdot \mathbf{x} + b = 0$: when training data $(\mathbf{x}_i, y_i), i = 1,...,l$ are separated by this hyperplane, then $y_i(\mathbf{w} \cdot \mathbf{x}_i + b) \geq 1$, where $y_i = \pm 1$ are the labels. It can be shown that the margin is $2/\|\mathbf{w}\|$, so finding the hyperplane which separates data with maximal margin means to:

$$\begin{cases} \text{minimize } \|\mathbf{w}\|^2/2 \\ \text{with } y_i(\mathbf{w} \cdot \mathbf{x}_i + b) \geq 1. \end{cases} \qquad (1)$$

In order to allow for misclassication errors, constraints are relaxed to

$y_i(\mathbf{w} \cdot \mathbf{x}_i + b) \geq 1 - \xi_i, \ \xi_i \geq 0.$ (1) becomes then:

$$\begin{cases} \text{minimize } \|\mathbf{w}\|^2/2 + C \cdot \sum \xi_i \\ \text{with } y_i(\mathbf{w} \cdot \mathbf{x}_i + b) \geq 1 - \xi_i. \end{cases} \tag{2}$$

The dual formulation of (2) reduces to:

$$\begin{cases} \text{maximize } \sum \alpha_i - \dfrac{1}{2} \sum \alpha_i \alpha_j (\mathbf{x}_i \cdot \mathbf{x}_j) y_i y_j \\ \quad \text{with } \sum \alpha_i y_i = 0, 0 \leq \alpha_i \leq C. \end{cases} \tag{3}$$

In this formulation example vectors are present only in dot products and the execution of

point a) becomes quite simple, because of a theorem by Mercer (Vapnik 1995). It gives an

easy way to compute dot products in features space, where vectors of input space are non-

linearly mapped by a function $\phi(\mathbf{x})$.

By using a suitable function K such that $\phi(\mathbf{x}_i) \cdot \phi(\mathbf{x}_j) = K(\mathbf{x}_i, \mathbf{x}_j)$ we do not need to calculate

each singular mapping $\phi(\mathbf{x})$. For our purposes we need positive misclassified examples to

outweigh negative ones: in order to do this it is necessary to modify (2) in the following way:

$$\begin{cases} \text{minimize } \|\mathbf{w}\|^2/2 + C^+ \cdot \sum \xi_i^+ + C^- \cdot \sum \xi_i^- \\ \text{with } (\mathbf{w} \cdot \mathbf{x}_i + b) \geq 1 - \xi_i^+, (\mathbf{w} \cdot \mathbf{x}_i + b) \leq -1 + \xi_i^- \end{cases} \tag{4}$$

where C^+ and C^- give different costs to false-positive and false-negative errors.

We have chosen a polynomial *kernel*: $(\mathbf{x} \cdot \mathbf{y} + 1)^d$, with the d degree ranging from 2 to 7. For

each value of d we have constructed a Free Response Operating Characteristics (FROC)

curve, by changing the ratio C^+/C^-. It turned out that this ratio is much more important for fine

tuning than the choice of d or of the absolute value of C^-. We train the machine with *fpr*

features of signals coming from 20 images of the Nijmegen database.

3. Results

The goal of this paper is to show that the combination of different methods with an SVM

classifier improves the performance of the detection algorithm. In figure 4 three FROC's

curves with SVM classifier are depicted: the first related to the combined method, the second

for the coarse method and the third for the fine one. There is a clear improvement due to the

simultaneous use of the two methods. We yield a sensitivity of 95% with 0.6 false per image.

In figure 5 we compare our combined method with SVM with the same detection scheme

with a Multi-Layer Perceptron (MLP) classifier. Results with MLP are comparable to others

obtained on the same database (Karssemeijer 1993, Veldkamp and Karssemeijer 1998,

Brown et al. 1998, Mossi and Albiol 1998). We stress that the SVM classifier clearly
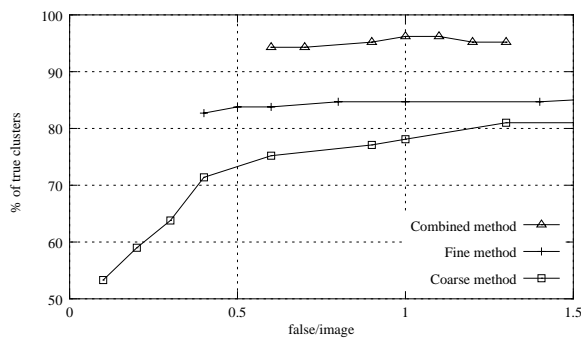
improves the performance of the detection scheme.

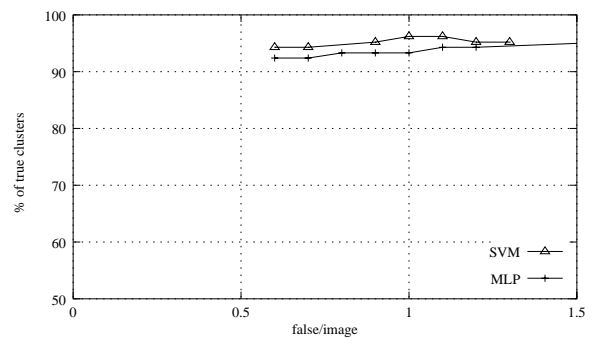**Figure 4.** FROC's of our detection system with SVM classifier.

Figure 5. FROC's of our detection system with SVM and MLP classifiers.

Breast Cancer Screening and of the Department of Radiology at the University of Nijmegen, the Netherlands.

References

Brown S., R. Li, L. Brandt, L. Wilson, G. Kossof, M. Kossof. 1998. Development of a multi-feature CAD system for mammography. Digital Mammography: Nijmegen, Kluwer Academic Pub. 189-196.

Chan H. P., K. Doi, S. Galhotra, C. J. Vyborny, H. MacMahon, P. M. Jokich. 1987. Image feature analysis and Computer Aided Diagnosis in digital radiography: automated detection of microcalcifications in mammography. Med. Phys. 14. 538-548.

Ema T., K. Doi, R. M. Nishikawa, Y. Jiang, J. Papaioannou. 1995. Image feature analysis and Computer Aided Diagnosis in digital radiography: reduction of false-positive clustered microcalcifications using local edge-gradient analysis. Med. Phys. 22. 161-169.

Gürcan M. N., Y. Yardimci, A. E. Cetin. 1998. Microcalcifications detection using adaptive filtering and gaussianity tests. Digital Mammography: Nijmegen, Kluwer Academic Pub. 157-164.

Karssemeijer N. 1993. Adaptive noise equalization and recognition of micro-calcification clusters in mammograms. International Journal of Pattern Recognition and Artificial Intelligence 7. 1357-1376.

Lado M. J., P. G. Tahoces, A. J. Mendez, M. Souto, J. J.Vidal. 1999. A wavelet based algorithm for detecting clustered microcalcifications in digital mammograms. Med. Phys. 26. 1294-1305.

Mossi J. M., A. Albiol. 1998. Automatic detection of clustered microcalcifications using morphological reconstruction. Digital Mammography: Nijmegen, Kluwer Academic Pub. 475-476.

Yoshida H., K. Doi, R. M. Nishikawa. 1994. Automatic detection of clustered microcalcifications in digital mammograms using wavelet transform techniques. Proc. SPIE 2167. 868-886.

Yoshida H., K. Doi, R. M. Nishikawa, M. L. Giger, R. A. Schmidt. 1996. An improved CAD scheme using wavelet transform for detection of clustered microcalcifications in digital mammograms. Acad. Radiol. 3. 621-627.

Vapnik V. 1995. The nature of statistical learning theory. Springer Verlag.

Veldkamp W., N. Karssemeijer. 1998. Improved correction for signal dependent noise applied to automatic detection of microcalcifications. Digital Mammography: Nijmegen, Kluwer Academic Pub. 169-176.